

Grundlagen der Künstlichen Intelligenz

1. Einführung: Was ist Künstliche Intelligenz?

Malte Helmert

Universität Basel

20. Februar 2015

Grundlagen der Künstlichen Intelligenz

20. Februar 2015 — 1. Einführung: Was ist Künstliche Intelligenz?

1.1 Was ist KI?

1.2 Menschlich handeln

1.3 Menschlich denken

1.4 Rational denken

1.5 Rational handeln

1.6 Zusammenfassung

Einführung: Überblick

Kapitelüberblick Einführung:

- ▶ 1. Was ist Künstliche Intelligenz?
- ▶ 2. KI früher und heute
- ▶ 3. Rationale Agenten
- ▶ 4. Umgebungen und Problemlösungsverfahren

1. Einführung: Was ist Künstliche Intelligenz?

Was ist KI?

1.1 Was ist KI?

Was ist KI?

Was versteht man unter **künstlicher Intelligenz**?

↪ keine allgemein akzeptierte Definition!

Oft pragmatische Definitionen:

- ▶ „KI ist, was Kleriker machen.“
- ▶ „KI ist das Lösen von schweren Problemen.“

Im Folgenden: Vorstellung einiger verbreiteter Ideen

Was ist KI: menschlich vs. rational; denken vs. handeln

„[the automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning“ (Bellman, 1978)	„the study of mental faculties through the use of computational models“ (Charniak & McDermott, 1985)
„the study of how to make computers do things at which, at the moment, people are better“ (Rich & Knight, 1991)	„the branch of computer science that is concerned with the automation of intelligent behavior“ (Luger & Stubblefield, 1993)

Vier typische Kategorien:

menschlich denken	rational denken
menschlich handeln	rational handeln

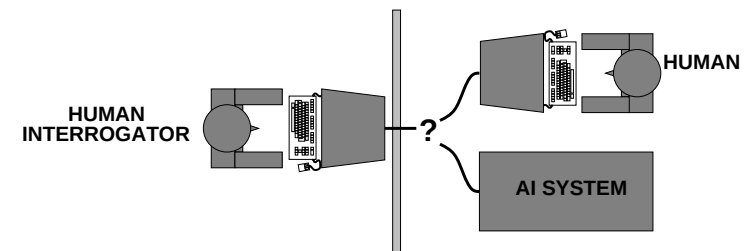
↪ hier (und heutzutage vorherrschend): rational handeln

1.2 Menschlich handeln

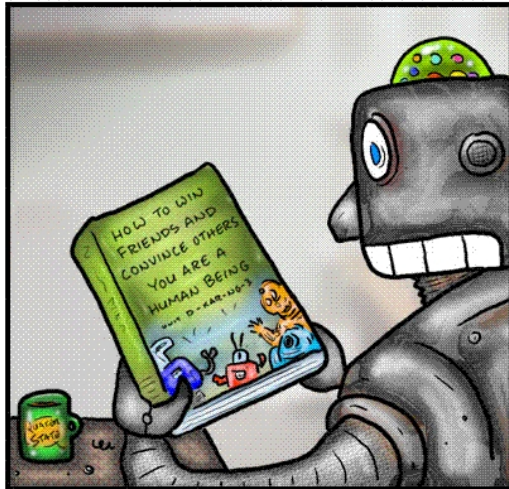
Menschlich handeln: der Turing-Test

Alan Turing, *Computing Machinery and Intelligence* (1950):

- ▶ Von „Können Maschinen denken?“ zu „Können Maschinen sich intelligent verhalten?“
- ▶ Operationalisierung: das **Imitationsspiel**



Cartoon



Unit Bob crams for his Turing Test.

Wie nützlich ist der Turing-Test?

Turing-Test:

- ▶ Wissenschaftlichkeit ist fragwürdig
- ▶ spielt im KI-„Mainstream“ keine Rolle mehr
- ▶ dennoch: jährliche Wettbewerbe (Loebner Prize):
<http://www.loebner.net/Prizef/loebner-prize.html>
- ▶ praktische Anwendung: **CAPTCHA** („Completely Automated Public Turing test to tell Computers and Humans Apart“)



Turing und der Turing-Test im Kino



<http://www.imdb.com/title/tt2084970/>

Turings „Computing Machinery and Intelligence“

Turings **Computing Machinery and Intelligence**:

- ▶ diskutierte bereits alle wesentlichen Argumente des 20. Jahrhunderts gegen Möglichkeit von KI
- ▶ schlug Kernaspekte der KI vor: Wissensrepräsentation, Schlussfolgern, Sprachverständnis, **Lernen**
- ▶ Vorhersage: im Jahr 2000 kann eine Maschine mit 30%iger Wahrscheinlichkeit Laien für 5 Minuten zum Narren halten
- ▶ kürzlich in den Nachrichten:
<http://www.engadget.com/2014/06/08/supercomputer-passes-turing-test/>

1.3 Menschlich denken

Menschlich denken: Kognitionswissenschaft

- ▶ **Kognitive Revolution** der 1960er: Informationsverarbeitung ersetzt in der Psychologie den dominierenden Behaviorismus
- ▶ Welche kognitiven Fähigkeiten sind für intelligentes Verhalten nötig?
- ▶ erfordert wissenschaftliche Theorien der Gehirnaktivität
 - ↪ Welches Abstraktionsniveau? „**Wissen**“ oder „**Schaltkreise**“?
- ▶ Wie überprüfen? Erfordert
 - ▶ Vorhersage/Test menschlichen Verhaltens (Top-Down) oder
 - ▶ Identifikation aus neurologischen Daten (Bottom-Up)
- ▶ entspricht grob **Kognitionswissenschaft** und **kognitiver Neurowissenschaft**
- ▶ heute von der KI getrennte Gebiete

1.4 Rational denken

Rational denken: Gesetze des Denkens

- ▶ **normativ (präskriptiv)** statt **deskriptiv**
- ▶ **Aristoteles**: Was sind korrekte Argumente/Denkprozesse?
- ▶ **Syllogismen**: Strukturen für Argumente, die immer korrekte Schlussfolgerungen liefern, sofern die Prämissen korrekt sind:
 - ▶ Sokrates ist ein Mensch.
 - ▶ Alle Menschen sind sterblich.
 - ▶ **Also** ist Sokrates sterblich.
- ▶ Mehrere griechische Schulen entwickelten verschiedene Formen von **Logik**:
 - ▶ **Notationen** (Syntax) und
 - ▶ **Ableitungsregeln** (Kalküle) für „richtiges“ Denken
- ▶ über die mathematische Logik (Anfang 20. Jh.) direkte Verbindung zur modernen KI

Probleme des logischen Ansatzes

Probleme:

- ▶ nicht alles intelligente Verhalten entspringt logischem Denken
- ▶ Welche Schlussfolgerungen sind **relevant**?
- ▶ Wie mit **Unsicherheit** umgehen?
- ▶ Wie mit **Widersprüchen** umgehen?

1.5 Rational handeln

Rational handeln

Rationales Verhalten: „das Richtige tun“

- ▶ das Richtige: **maximiere Nutzen**
gegeben **verfügbare Information**
- ▶ erfordert nicht zwangsläufig „Denken“ (z. B. Reflexe)

KI als **Entwurf rationaler Agenten** hat zwei Vorteile:

- ▶ **allgemeiner** als „Gesetze des Denkens“:
 ↪ logische Inferenz nur **ein** Mechanismus
 zum Erreichen rationalen Verhaltens
- ▶ besser der **wissenschaftlichen Methode** zugänglich
 als Ansätze, die auf menschlichem Verhalten/Denken fussen

1.6 Zusammenfassung

Zusammenfassung

Was ist KI? $\rightsquigarrow$ mehrere mögliche Antworten

- ▶ Orientierung am **Menschen** vs. am Nutzen (**Rationalität**)
- ▶ Geht es um das äussere **Handeln** oder das innere **Denken**?

$\rightsquigarrow$ vier Kombinationen:

- ▶ menschlich handeln: Beispiel Turing-Test
- ▶ menschlich denken: vgl. Kognitionswissenschaft
- ▶ rational denken: Logik
- ▶ rational handeln: die heute häufigste Sichtweise
 $\rightsquigarrow$ ermögliche wissenschaftliche Methode